MEMORANDUM

DEPARTMENT OF ENVIRONMENTAL QUALITY
DIVISION OF WATER PROGRAMS COORDINATION
OFFICE OF SPILL RESPONSE AND REMEDIATION

Mail Address:	Location:
P.O. Box 10009	629 East Main Street
Richmond, VA 23240	Richmond, VA 23219

SUBJECT: Guidance Memorandum Number 02-2013
Evaluation of Vapor Monitoring Data for Petroleum Storage Tank Release Detection

TO: Regional Directors

FROM: Larry Lawson, P.E., Director

DATE: July 17, 2002

COPIES: Regional Ground Water Managers, J. Andrew Hagelin, Fred Cunningham, Robyne Bridgman, Russ Ellison, Sam Lillard, Dave Chance

Summary

Effective the date of this memorandum, staff should use this guidance to evaluate vapor monitoring data collected as part of release detection at UST and AST facilities. This guidance provides a procedure that staff and storage tank owners/operators may use to evaluate vapor monitoring data to determine if a release should be suspected. Included with this guidance is a spreadsheet entitled: *Statistics-Vapor Monitoring for Release Detection* that performs the statistical calculations discussed in this guidance document.

This guidance will be incorporated into the next editions of the Storage Tank Program Compliance Manual and the Storage Tank Program Technical Manual.

Electronic copies of this guidance are in PDF and Excel format and may be read online, downloaded, and distributed to DEQ staff and/or the public. The numbering convention used for guidance memoranda is: Guidance Memorandum, then a two digit number designating the year in which the memorandum was issued, followed by a hyphen and the document number.

Contact Information

Should you have any questions or need further assistance on this guidance, please contact:
James Barnett (804) 698-4289
jsbarnett@deq.state.va.us

Russ Ellison (804) 698-4269
rpellison@deq.state.va.us

Disclaimer

This document provides procedural guidance to the DEQ Storage Tank Program staff. This document is guidance only. It does not establish or affect legal rights or obligations. It does not establish a binding norm and is not finally determinative of the issues addressed. Agency decisions in any particular case will be made by applying the State Water Control Law and the implementation regulations on the basis of site-specific facts.

Evaluation of Vapor Monitoring Data for Release Detection

Issue Statement

Vapor monitoring is one of the options that tank owners/operators have to meet the release detection requirements of the UST Technical Regulation or the Facility and AST Regulation. Vapor monitors used for release detection must be able to detect any significant increase in concentration above the background of the regulated substance stored in the tank system. The issue regarding what constitutes a "significant" increase in vapor concentrations at a site has been raised by DEQ staff that perform AST and UST compliance inspections. This memo provides guidelines and a procedure that staff and tank owners/operators may use to evaluate vapor monitoring data to determine if a release should be suspected.

Regulatory Requirements - Vapor Monitoring for Release Detection

Release detection is required for regulated USTs and ASTs. Section 160 of the UST Technical Regulation requires vapor monitoring systems be designed to detect releases at the earliest time possible. The Facility and AST Regulation requires the system of release detection that will be used at the facility to be described in the Oil Discharge Contingency Plan submitted for that facility.

Vapor monitoring is one of the release detection options at storage tank facilities. Vapor monitoring systems must be placed in the backfill around the tank(s) and the backfill must be sufficiently porous to allow migration of vapors from the tank system to the vapor monitoring points. The regulated substance or tracer stored in the tank must be sufficiently volatile to be detected if it is released. The measurement of vapors cannot be rendered inoperative by ground water changes, rainfall, soil moisture, or other site conditions. The level of background contamination at the site must not interfere the detection of releases and the site must be assessed to ensure compliance with regulatory requirements. Tanks and lines must be monitored for releases at least once every thirty (30) days. The monitoring method used also should not allow a release to go undetected for greater than 30 days.

Recommendations for Determining if a Release Should be Suspected

In order to comply with the requirement of timely release discovery, tank owners/operators must evaluate the site to determine if vapor monitoring is an appropriate release detection method for the site. The method used to determine if a release has occurred should account for site conditions including background vapor concentrations and the type of vapor monitoring equipment used at the site. Vapor monitoring equipment may be separated into two primary categories: (1) vapor sensors that are part of an automated tank monitoring system; and (2) vapor monitors that require the user to evaluate the data produced by the monitor to determine the significance of that data.

Vapor Sensors that are Part of an Automated Tank Monitoring System

Vendors of storage tank equipment have developed sensors that monitor vapor concentrations in vapor monitoring wells and provide this information directly to an automated tank monitoring system. These automated vapor monitors have pre-set vapor thresholds that alert the control unit for the tank system if vapor concentrations are above the factory set thresholds. Tank owners/operators using this type of vapor monitoring device need to demonstrate that the background conditions at their site will not interfere with the operation of the vapor sensor(s). After evaluating the site and determining that background conditions will not interfere with the proper operation of the vapor sensor, the tank owner/operator may use the pre-set threshold on the vapor sensor as the point at which a release should be suspected.

Vapor Monitoring Devices Requiring Action by User

Vapor monitoring devices that are not part of an automated system require the tank owner/operator to evaluate the monitoring data, determine the significance of that data, and determine if a release should be suspected. In order to determine the significance of monthly vapor monitoring data, tank owners/operators must determine background conditions at the site. Monthly monitoring data then must be compared with the background data to determine if a release should be suspected.

NOTE: Consistency in the vapor sample collection procedures and equipment used is of great importance. Changing instruments (e.g. PID to FID, FID to detector tube) will make it difficult to compare vapor readings and, ultimately, make it difficult to determine the significance of monthly monitoring data.

Determination of Background Vapor Conditions at a Site

Tank owners/operators using devices that require the operator of the device to determine the significance of the data, need to determine background vapor concentrations in each vapor monitoring well at the site. At sites with new vapor monitoring systems, background vapor concentrations must be determined within the first month after the AST or UST system is installed and filled with product. Background vapor concentrations must be determined for each vapor monitoring well by taking at least 20 vapor measurements from each well during the first month of operation. This background data must be collected using the same equipment and procedures that will be used to collect all subsequent vapor data. The data set for each well then may be considered to represent background conditions in that well.

Background vapor concentrations also must be determined for each vapor monitoring point at sites where the tank owner/operator is already using vapor monitoring as a release detection method and has never evaluated background conditions. Provided that a release has not occurred since the vapor monitoring system was installed, tank owners/operators at sites that are already using vapor monitoring as a release detection method may use their existing monthly vapor monitoring data to determine background vapor concentrations. As with the new installations, at least 20 vapor monitoring readings for each vapor monitoring well are needed to derive a background vapor concentration for that well. If the tank owner/operator has fewer than 20 vapor monitoring observations for a particular well, additional vapor readings must be taken during the first month after the effective date of this guidance so that there are at least 20 vapor monitoring observations for the well. For example, a tank owner/operator having 12 months of vapor monitoring data for a particular vapor monitoring well will need to collect at least 8 vapor readings during the month following the effective date of this guidance to determine the background vapor concentration for that well. The tank owner/operator then must evaluate the background data set and determine the concentration at which a release should be suspected.

NOTES:

1. Petroleum vapors are heavier than the major gases in the atmosphere and will tend to be found in greatest concentration at or near the bottom of a vapor monitoring well. Sample collection procedures used at a site must account for this physical characteristic.
2. If the upper tolerance limit for the background data exceeds the upper range on the vapor monitoring instrument, vapor monitoring cannot be used as a release detection method at the site.

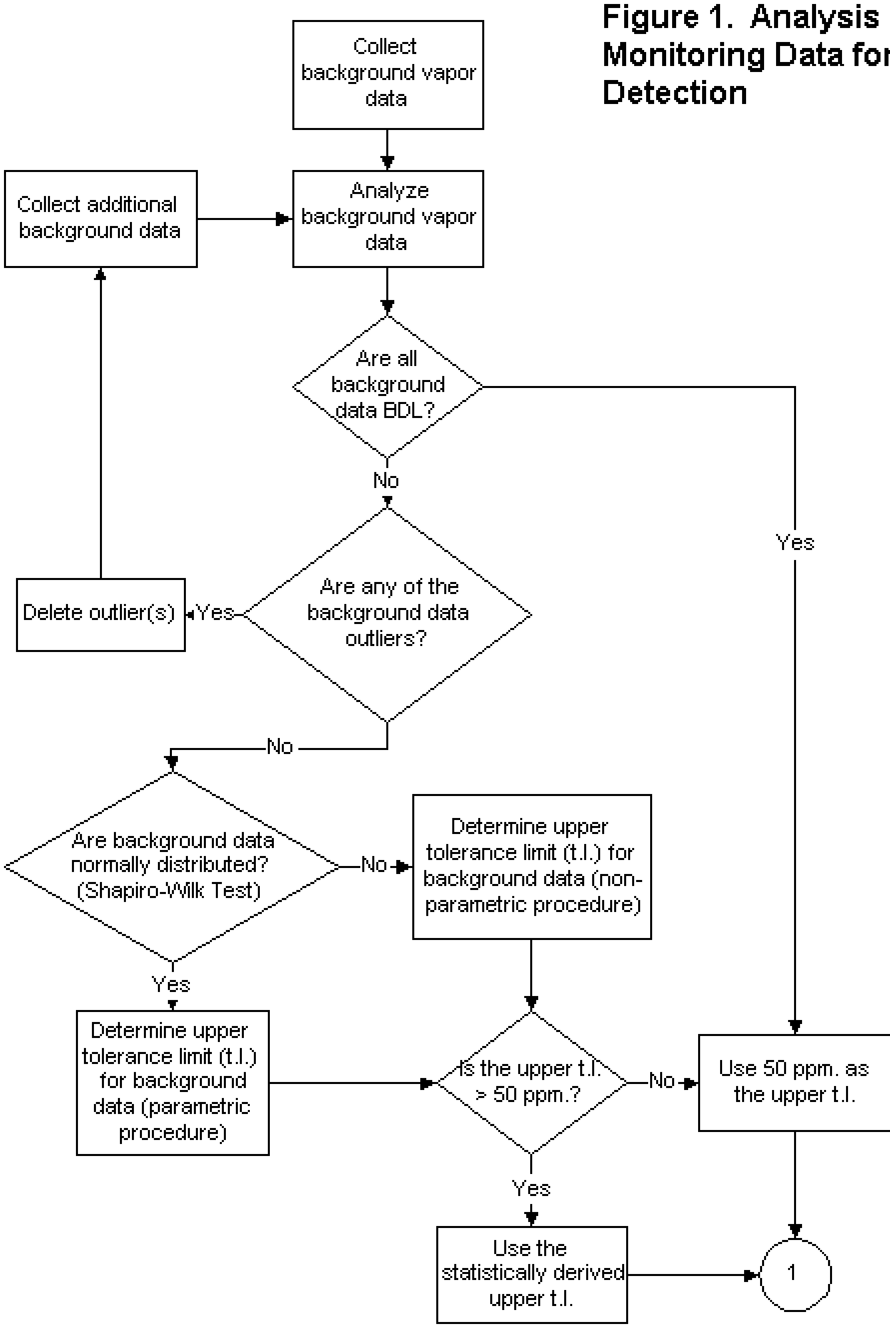
If no releases occur during the first year of system operation, the tank owner/operator may re-calculate the background concentration for each well by using the original background data and the monthly monitoring data from the previous year. This additional data should provide an even greater amount of information about ambient conditions (i.e. background) at the site. At the end of subsequent years, tank owners/operators may use data from that year and combine that data with monitoring data from the previous one or two years in determining background conditions at the site provided that no releases occurred during the year.

Data Evaluation Procedure

In order to determine if monthly monitoring data indicate a release may have occurred, the tank owner/operator, consultant, or DEQ staff needs to establish an upper tolerance limit for the background data. If the vapor concentration in the monthly monitoring sample for a well exceeds the upper tolerance limit, a release may have occurred. If the upper tolerance limit is exceeded, the tank owner/operator should report a suspected release to DEQ and collect at least four additional vapor samples from that well over the next five days. The mean of the monthly result that triggered the investigation plus the additional investigative samples then may be compared with the mean of the background data. If the mean of the monthly result plus investigative samples is statistically greater than the background mean, there is reason to suspect a release and the tank owner/operator should determine if the tank is leaking. See Figure 1 for a flowchart explaining a procedure that may be used for determining if a release should be suspected from vapor monitoring data.

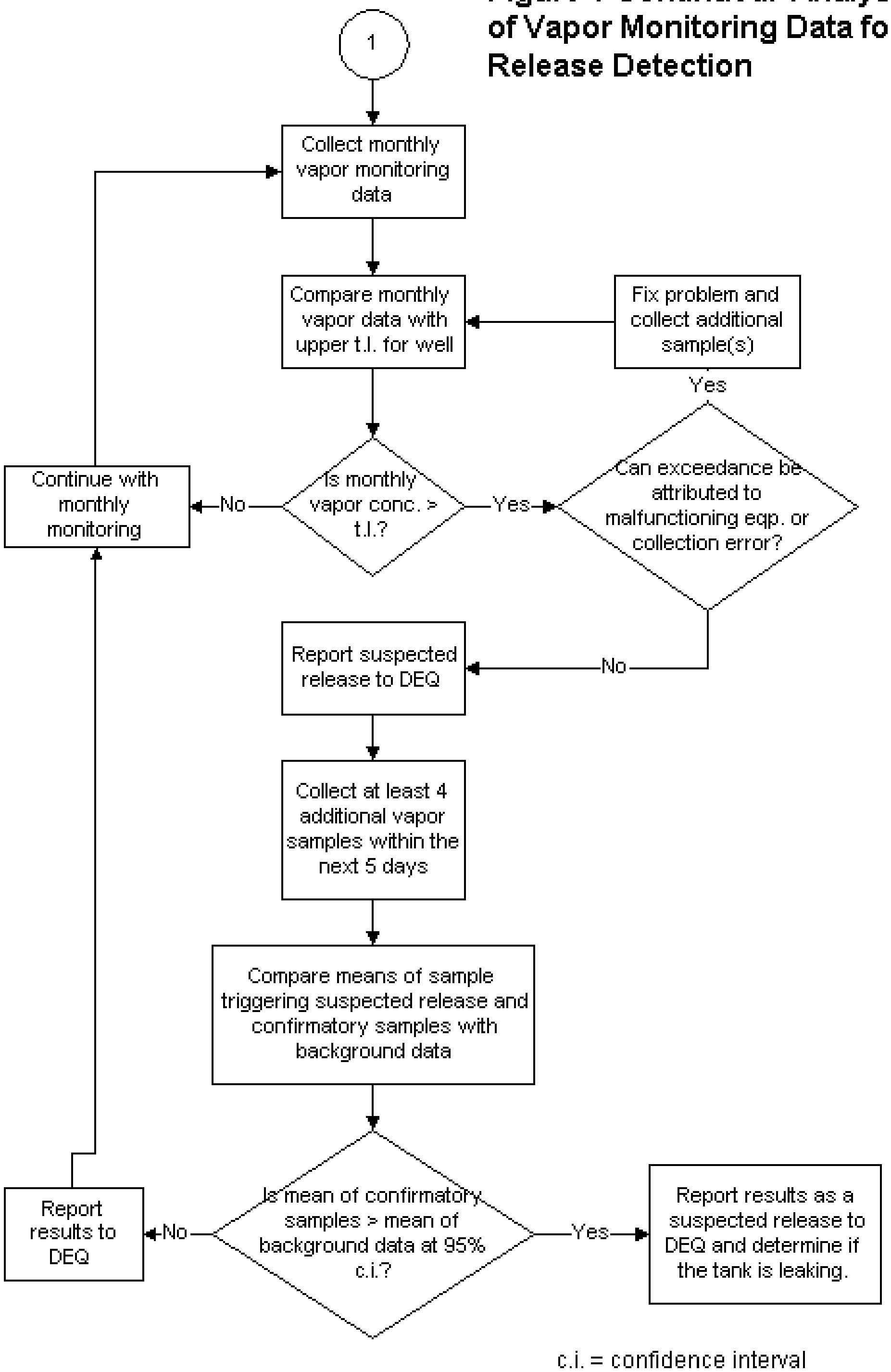
Statistical methods may be used to evaluate vapor monitoring data, estimate "background" concentrations, and determine when the data indicate a release should be suspected. Staff, tank owners/operators, and consultants may use the statistical procedure that follows to evaluate vapor monitoring data. A spreadsheet that will perform these statistical calculations is available for use by staff and tank owners/operators and is entitled: *Statistics-Vapor Monitoring for Release Detection*. A User's Guide explaining how to use the spreadsheet is attached to this guidance document.

Figure 1. Analysis of Vapor Monitoring Data for Release Detection



BDL = below detection limit

Figure 1 Continued. Analysis of Vapor Monitoring Data for Release Detection



1. Analyze Background Data

After the background data are collected, the data must be analyzed to determine if any of the observations is an outlier. Outliers are data values that are unusually high or low relative to the rest of the data. Often, outliers may be attributed to sample collection, laboratory, or data entry errors. Data values that are outliers must be removed from the background data set in order to preserve the integrity of the data and allow a more accurate estimate of "background conditions" to be obtained.

A procedure that may be used to determine if data points are outliers is as follows:

1. Identify the potential outlier

2. Using all data points, including the suspected outlier, compute the mean and the standard deviation of the data set.

(1) Mean $$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

where: n = number of samples (vapor concentrations)
X_i = value of the i^{th} sample

(2) Sample Variance: $$s^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n - 1}$$

(3) Sample Standard Deviation: $s = (s^2)^{1/2}$

3. Compute the outlier test statistic (T_n)

(4) $$T_n = (X_i - \bar{X})/s$$

4. Compare the outlier test statistic T_n with the critical point (t_c) from Table 1. If $T_n > t_c$, there is evidence that the suspected outlier is not consistent with the pattern of the data.

NOTE: A minimum of 20 data points are needed to estimate "background" conditions at a site. Outliers must be removed from the data background data set and additional background data will need to be collected if the removal of outliers results in less than 20 background data values.

Table 1. a - Level Critical Points (t_c) for One-Tailed Outlier Test					
n	t_c	n	t_c	n	t_c
20	2.557	30	2.759	41	2.877
21	2.580	31	2.773	42	2.887
22	2.603	32	2.783	43	2.896
23	2.624	33	2.786	44	2.905
24	2.644	34	2.799	45	2.914
25	2.663	35	2.811	46	2.923
26	2.681	36	2.823	47	2.931
27	2.698	37	2.835	48	2.940
28	2.714	38	2.846	49	2.948
29	2.730	39	2.857	50	2.956
30	2.745	40	2.866		

2. Evaluate Data Distribution

The distribution of the background data set determines the type(s) of statistical tests that may be used to analyze the data. Data that are normally distributed may be analyzed using parametric tests such as those based on the Student's t test. Data that are not normally distributed may be analyzed using non-parametric techniques.

NOTE: If all background data observations are below the instrument detection limit, it is not necessary to evaluate the data distribution. Non-parametric statistical techniques may be used to further analyze the data.

The Shapiro-Wilk test may be used to evaluate the background data set and determine if the data set is normally distributed. The Shapiro-Wilk test (Example 1) compares the quantiles from the sample data set (in this case, the background vapor samples) to the corresponding quantiles from a normal distribution. This test assumes that if the data set is normally distributed, the quantiles from that data set will be highly correlated with the quantiles from a normal distribution. The procedure for performing the Shapiro-Wilk test is as follows:

1. Order the data set from the smallest value to the greatest.

2. Compute the differences $[X_{(n-i+1)} + X_{(i)}]$ for each i = 1...n.

   X = data point (i.e. vapor concentration)
   n = number of samples

3. Determine k where k = n/2
   If n/2 is not an integer, round to the next highest whole number (e.g. if n/2 is 20.5, use 21)

4. Determine the Shapiro-Wilk coefficient a_{n-i+1} for i = 1...k. Use Table 2 (EPA 1998) to determine the Shapiro-Wilk Coefficient.

5. Compute b where:

$$(5) \qquad b = \sum_{i=1}^{k} b_i = \sum_{i=1}^{k} a_{n-i+1} (X_{(n-i+1)} X_{(i)})$$

6. Compute the Shapiro-Wilk test statistic

$$(6) \qquad W = [b/s(n-1)^{1/2}]^2$$

7. Determine the critical point of the test for a selected a using Table 3 (EPA 1998). Compare the Shapiro-Wilk Statistic (W) to the critical point. If the Shapiro-Wilk Statistic for the data set exceeds the critical point, there is no evidence to reject the assumption that the data are normally distributed. If the Shapiro-Wilk statistic is less than the critical point, there is evidence at the a level of significance that the background data are not normally distributed. If the background data are not normally distributed, non-parametric techniques should be used to evaluate the data.

3. Establish a Tolerance Limit for the Background Data

Monthly vapor monitoring measurements taken after background concentrations are established may be compared with a tolerance (TL) limit established for the background data set. A tolerance limit is a limit with a specified degree of coverage and level of confidence. A tolerance limit having 95 percent coverage and 95 percent confidence can be interpreted to contain at least 95 percent of the distribution of observations from background data with 95 percent confidence (DEQ 1998).

The Storage Tank Program will use a tolerance limit of 95% (i.e. this interval is expected to contain at least 95% of the distribution of observations from the background dataset with a 95% confidence). The upper 95th tolerance limit is justified by its consistency with the requirements for other release detection methods specified by the UST Technical Regulation.

If the statistically calculated upper tolerance limit is below 50 ppm., the Storage Tank Program will use a total organic vapor concentration of 50 ppm. as an upper tolerance limit for the concentration at which additional data needs to be collected to determine if a release should be suspected. According to studies performed by the National Work Group on Leak Detection Evaluations, the detection limits for most of the automated vapor detectors presently on the market and tested by this group range from around 50 ppm. to 500 ppm. The DEQ believes that using 50 ppm. of total organic vapors as an upper tolerance limit in those cases where the calculated upper tolerance limit is below 50 ppm.: (1) is consistent with the automated leak detectors on the market; and (2) will reduce the number of suspected releases caused by false positive vapor monitoring results.

Table 2. Coefficients [a_{n-i+1}] for Shapiro-Wilk Test of Normality, n = 2(1)50 (EPA 1998)										
i/n	2	3	4	5	6	7	8	9	10	
1	.7071	.7071	.6872	.6646	.6431	.6233	.6052	.5888	.5739	
2		0	.1677	.2413	.2806	.3031	.3164	.3244	.3291	
3				0	.0875	.1401	.1743	.1976	.2141	
4						0	.0561	.0947	.1224	
5								0	.0399	
i/n	11	12	13	14	15	16	17	18	19	20
1	.5601	.5475	.5359	.5251	.5150	.5056	.4968	.4886	.4808	.4734
2	.3315	.3325	.3325	.3314	.3306	.3290	.3273	.3253	.3232	.3211
3	.2260	.2347	.2412	.2460	.2495	.2521	.2543	.2553	.2561	.2565
4	.1429	.1586	.1707	.1802	.1878	.1939	.1988	.2027	.2059	.2085
5	.0695	.0922	.1099	.1240	.1353	.1447	.1524	.1587	.1641	.1686
6	0	.0303	.0539	.0727	.0880	.1005	.1109	.1197	.1271	.1334
7			0	.0240	.0433	.0593	.0725	.0837	.0932	.1013
8					0	.0196	.0359	.0496	.0612	.0711
9							0	.0163	.0303	.0422
10									0	.0140
i/n	21	22	23	24	25	26	27	28	29	30
1	.4643	.4590	.4542	.4493	.4450	.4407	.4366	.4328	.4291	.4254
2	.3185	.3156	.3126	.3098	.3069	.3043	.3018	.2992	.2968	.2944
3	.2578	.2571	.2563	.2554	.2543	.2533	.2522	.2510	.2499	.2487
4	.2119	.2131	.2139	.2145	.2148	.2151	.2152	.2151	.2150	.2148
5	.1736	.1764	.1787	.1807	.1822	.1836	.1848	.1857	.1864	.1870
6	.1399	.1443	.1480	.1512	.1539	.1563	.1584	.1601	.1616	.1630
7	.1092	.1150	.1201	.1245	.1283	.1316	.1346	.1372	.1395	.1415
8	.0804	.0878	.0941	.0997	.1046	.1089	.1128	.1162	.1192	.1219
9	.0530	.0618	.0696	..0764	.0823	.0876	.0923	.0965	.1002	.1036
10	.0263	.0368	.0459	.0539	.0610	.0672	.0728	.0778	.0822	.0862
11	0	.0122	.0228	.0321	.0403	.0476	.0430	.0598	.0650	.0697
12			0	.0107	.0200	.0284	.0358	.0424	.0483	.0537
13					0	.0094	.0178	.0253	.0320	.0381
14							0	.0084	.0159	.0227
15									0	.0076

Table 2 Continued. Coefficients $[a_{n-i+1}]$ for Shapiro-Wilk Test of Normality, n = 2(1)50 (EPA 1998)

i/n	31	32	33	34	35	36	37	38	39	40
1	.4220	.4188	.4156	.4127	.4096	.4068	.4040	.4015	.3989	.3964
2	.2921	.2898	.2876	.2854	.2834	.2813	.2794	.2774	.2755	.2737
3	.2475	.2463	.2451	.2439	.2427	.2415	.2403	.2391	.2380	.2368
4	.2145	.2141	.2137	.2132	.2127	.2121	.2116	.2110	.2104	.2098
5	.1874	.1878	.1880	.1882	.1883	.1883	.1883	.1881	.1880	.1878
6	.1641	.1651	.1660	.1667	.1673	.1678	.1683	.1686	.1689	.1691
7	.1433	.1449	.1463	.1475	.1487	.1496	.1503	.1513	.1520	.1526
8	.1243	.1265	.1284	.1301	.1317	.1331	.1344	.1356	.1366	.1376
9	.1066	.1093	.1118	.1140	.1160	.1179	.1196	.1211	.1225	.1237
10	.0899	.0931	.0961	.0988	.1013	.1036	.1056	.1075	.1092	.1108
11	.0739	.0777	.0812	.0844	.0873	.0900	.0924	.0947	.0967	.0986
12	.0585	.0629	.0669	.0706	.0739	.0770	.0798	.0824	.0848	.0870
13	.0435	.0485	.0530	.0572	.0610	.0645	.0677	.0706	.0733	.0759
14	.0289	.0344	.0395	.0441	.0484	.0523	.0559	.0592	.0622	.0651
15	.0144	.0206	.0262	.0314	.0361	.0404	.0444	.0481	.0515	.0546
16	0	.0068	.0131	.0187	.0239	.0287	.0331	.0372	.0409	.0444
17			0	.0062	.0119	.0172	.0220	.0264	.0305	.0343
18					0	.0057	.0110	.0158	.0203	.0244
19							0	.0053	.0101	.0146
20									0	.0049
i/n	41	42	43	44	45	46	47	48	49	50
1	.3940	.3917	.3894	.3872	.3850	.3830	.3808	.3789	.3770	.3751
2	.2719	.2701	.2684	.2667	.2651	.2635	.2620	.2604	.2589	.2574
3	.2357	.2345	.2334	.2323	.2313	.2302	.2291	.2281	.2271	.2260
4	.2091	.2085	.2078	.2072	.2065	.2058	.2052	.2045	.2038	.2032
5	.1876	.1874	.1871	.1868	.1865	.1862	.1859	.1855	.1851	.1847
6	.1693	.1694	.1695	.1695	.1695	.1695	.1695	.1693	.1692	.1691
7	.1531	.1535	.1539	.1542	.1545	.1548	.1550	.1551	.1553	.1554
8	.1384	.1392	.1398	.1405	.1410	.1415	.1420	.1423	.1427	.1430
9	.1249	.1259	.1269	.1278	.1286	.1293	.1300	.1306	.1312	.1317
10	.1123	.1136	.1149	.1160	.1170	.1180	.1189	.1197	.1205	.1212
11	.1004	.1020	.1035	.1049	.1062	.1073	.1085	.1095	.1105	.1113
12	.0891	.0909	.0927	.0943	.0959	.0972	.0986	.0998	.1010	.1020
13	.0782	.0804	.0824	.0842	.0860	.0876	.0892	.0906	.0919	.0932
14	.0677	.0701	.0724	.0745	.0775	.0785	.0801	.0817	.0832	.0846
15	.0575	.0602	.0628	.0651	.0673	.0694	.0713	.0731	.0748	.0764
16	.0476	.0506	.0534	.0560	.0584	.0607	.0628	.0648	.0667	.0685
17	.0379	.0411	.0442	.0471	.0479	.0522	.0546	.0568	.0588	.0608
18	.0283	.0318	.0352	.0383	.0412	.0439	.0465	.0489	.0511	.0532
19	.0188	.0227	.0263	.0296	.0328	.0357	.0385	.0411	.0436	.0459
20	.0094	.0136	.0175	.0211	.0245	.0277	.0307	.0335	.0361	.0386
21	0	.0045	.0087	.0126	.0163	.0197	.0229	.0259	.0288	.0314
22			0	.0042	.0081	.0118	.0153	.0185	.0215	.0244
23					0	.0039	.0076	.0111	.0143	.0174
24							0	.0037	.0071	.0104
25									0	.0035

Table 3. a - Level Critical Points for Shapiro - Wilk Test, n = 3 (1) 50					
N/ a	.05 level significance	n/ a	.05 level significance	n/ a	.05 level significance
1	---	18	.897	35	.934
2	---	19	.901	36	.935
3	.767	20	.905	37	.936
4	.748	21	.908	38	.938
5	.762	22	.911	39	.939
6	.788	23	.914	40	.940
7	.803	24	.916	41	.941
8	.818	25	.918	42	.942
9	.829	26	.920	43	.943
10	.842	27	.923	44	.944
11	.850	28	.924	45	.945
12	.859	29	.926	46	.945
13	.866	30	.927	47	.946
14	.874	31	.929	48	.947
15	.881	32	.930	49	.947
16	.887	33	.931	50	.947
17	.892	34	.933		

Tolerance Limit - Normally Distributed Background Data

To calculate the upper 95th percent tolerance limit for the background data, determine the sample mean and standard deviation for the background data using equations 1 through 3. Next, use this mean and standard deviation along with the appropriate tolerance factor from Table 4 (Lieberman 1958) to determine the upper 95th percent tolerance limit for the background data set. Example 2 provides additional information about calculating an upper tolerance limit for normally distributed data.

(7) Sample Mean: $$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Where: X_i = the value of the ith sample
n = the number of samples

(8) $$TL = \bar{X} + k(n, ?, 1\text{-}a) * s$$

n = # of samples in background data set
k(n, ?, 1-a) = tolerance factor for one-sided normal tolerance interval

If the upper tolerance limit is less than 50 ppm., the tank owner/operator should use 50 ppm. as the upper tolerance limit.

Example 1. Shapiro-Wilk Test for Normality (DEQ 1998)

i	X_i	$X_{(n-i+1)}$	$X_{(n-i+1)} - X_i$	a_{n-i+1} (see Table 2)	b_i
1	1	942	941	.4734	445.47
2	3.1	637	633.9	.3211	203.55
3	8.7	578	569.3	.2565	146.03
4	10	331	321	.2085	66.93
5	14	292	248	.1686	41.81
6	19	151	132	.1334	17.61
7	21.4	85.6	64.2	.1013	6.5
8	27	81.5	54.5	.0711	3.87
9	39	64.4	25.4	.0422	1.07
10	56	58.8	2.8	.0140	.04
11	58.8	56	-2.8		
12	64.4	39	-25.4		
13	81.5	27	-54.5		
14	85.6	21.4	-64.2		
15	151	19	-132		
16	262	14	-248		
17	331	10	-321		
18	578	8.7	-569.3		
19	637	3.1	-633.9		
20	942	1	-941		

? b_i = 932.88

s = 259.72 (see equations 2 and 3)
W = $[932.88/259.72\ (19)^{1/2}]^2$
W = .679

The a = .05 level critical point for the Shapiro-Wilk test when n = 20 is .905. Since W < .905, there is evidence that the data do not follow a normal distribution..

NOTE: X_i is the background data set arranged in ascending order and $X_{(n-i+1)}$ is the background data set arranged in descending order.

Table 4. Tolerance Factors (k) for one-sided normal tolerance intervals with probability level (confidence factor) Y = .95 and coverage P = 95% (Lieberman, 1958)

N	k	n	k
3	7.655	20	2.396
4	5.145	21	2.371
5	4.202	22	2.350
6	3.707	23	2.329
7	3.399	24	2.309
8	3.188	25	2.292
9	3.031	30	2.220
10	2.911	35	2.166
11	2.815	40	2.125
12	2.736	45	2.092
13	2.670	50	2.065
14	2.614	55	2.036
15	2.566	60	2.017
16	2.523	65	2.000
17	2.486	70	1.986
18	2.443	75	1.972
19	2.423	100	1.924

Example 2. Establishing an upper tolerance limit for a normally distributed data set

Background data:

100	53	350	1000	1200	49	350	600	650	400
850	90	700	245	1345	75	130	550	920	240

$\bar{X}$ = 494.9
s = 401.6
k = 2.396 (from Table 3, 95% upper tolerance limit)

TL (tolerance limit) = 494.9 + 2.4 * 401.6
TL = 1459

Tolerance Limit - Background Data that are not Normally Distributed

When the background data are not normally distributed, the greatest concentration within the background data set may be used as an upper tolerance limit. The power of this type of tolerance limit, however, is heavily dependent upon the number of samples in the background data set. According to DEQ (1998), at least 60 background samples must be collected in order to guarantee at least 95 percent coverage.

In order to increase the coverage of the background data without having to collect 60 or more samples, an outlier test may be used to establish a concentration that would indicate the data point came from a population other than the background vapor conditions at the site. Individual data points within a sample of a population are evaluated using equation 4 to determine if any of the observations is an outlier. The potential outlier value (T_n) is then compared with a critical point for a one-tailed outlier test to determine if the observation is an outlier. Equation 4 may be re-arranged so that the concentration at which an observation is considered an outlier (X_n) may be determined (equation 9)

(9) $$X_n = T_n * s + \bar{X}$$

$T_n = t_c$, the critical point for the one-tailed outlier in Table 1.
n = # of background vapor observations

If the upper tolerance limit (in this case X_n) is less than 50 ppm., the tank owner/operator should use an upper tolerance limit of 50 ppm.

Example 3. Establishing an upper tolerance limit for a data set that is not normally distributed

Background data:

0	0	25	33	45	0	12	0	0	95
32	34	56	89	0	9	0	12	34	15

n = 20 $\bar{X}$ = 24.6 s = 28.7

$X_n = T_n * s + \bar{X}$
The critical value for Tn from Table 1 (n = 20) is 2.557

$X_n = 2.557 * 28.7 + 24.6$
$X_n = 98$

The upper tolerance limit that may be used for the background data set is 98.

4. Compare monthly data with the Upper Tolerance Limit

Once the upper 95th percent tolerance limit is estimated, the tank owner/operator should compare the monthly vapor reading for each well with the upper 95th percent tolerance limit for that well. If the monthly monitoring concentration is less than the upper tolerance limit for the well, the tank owner/operator should record the monthly monitoring concentration and continue on schedule with monthly monitoring. If a monthly monitoring reading exceeds the upper 95th percent tolerance limit for the vapor monitoring well, the tank owner/operator should: (1) report the results to DEQ; and (2) collect at least four additional vapor readings from that well over the next five days to determine if the vapor concentration in that well has increased significantly over "background" vapor concentrations in that well.

5. Compare the Means

After at least four additional vapor readings are taken from the well where the upper tolerance level was exceeded, the sample mean of these vapor readings (the monthly reading that triggered the investigation plus the additional readings) should be determined. This mean should be compared with the background mean for that well to determine if there has been a statistically significant increase in the vapor concentration in the well.

Comparison of Means - Normally Distributed Data

A test statistic that may be used to compare these two sample means when the data are normally distributed is Welch's t-test. Welch's t-test is calculated as follows:

(10) $$t = (\bar{X}_c - \bar{X}_b)/[\, s_c^2/n_c + s_b^2/n_b]^{1/2}$$

where: $\bar{X}_c$ = the mean concentration of the monthly and additional samples collected

$\bar{X}_b$ = the background mean
s_c^2 = variance of the monthly and additional samples
s_b^2 = background variance
n_c = number of samples (monthly and additional samples) collected
n_b = number of samples used to estimate the "background" concentration

The t-statistic derived by equation 10 should be compared with the critical t-statistic (t_c) for the degrees of freedom represented by the data. Equation 10 may be used to calculate the degrees of freedom for the data. After calculating the degrees of freedom, the critical t-value (t_c) may be looked up in the student's t-table (Table 4). See Example 4 for additional information about the use of Welch's t-test.

(11) $$df = [s_c^2/n_c + s_b^2/n_b]^2 \,/\, [(s_c^2/n_c)^2/(n_c-1) + (s_b^2/n_b)^2/(n_b-1)]$$

$$t_c = t_{(df,\ .95)}$$

If $t>t_c$, the null hypothesis of equal means between the two groups may be rejected and we may determine that the mean of the additional samples is significantly greater than the background mean.

Comparison of Means - Data that are not Normally Distributed

When additional data are collected following the exceedance of the upper tolerance limit, the data collected during the investigation need to be compared with the background data to determine if a release should be suspected. One way of performing this evaluation for data that are not normally distributed is to compare the upper confidence limit of the background mean with the mean of the monthly result that triggered the investigation plus the data collected as part of the investigation. The confidence limit is designed to contain the specified population parameter within a specified level of confidence or probability. For example, a 95 percent upper confidence limit of the mean having a value of 45 indicates that there is at least a 95 percent probability that the true mean of population is not greater than 45.

Table 4. Percentiles of the Student's t-Distribution

df/p	.95	df/p	.95	df/p	.95
1	6.314	23	1.714	45	1.679
2	2.920	24	1.711	46	1.679
3	2.353	25	1.708	47	1.678
4	2.132	26	1.706	48	1.677
5	2.015	27	1.703	49	1.677
6	1.943	28	1.701	50	1.676
7	1.895	29	1.699	51	1.675
8	1.860	30	1.697	52	1.675
9	1.833	31	1.696	53	1.674
10	1.812	32	1.694	54	1.674
11	1.796	33	1.692	55	1.673
12	1.782	34	1.691	56	1.673
13	1.771	35	1.690	57	1.672
14	1.761	36	1.688	58	1.672
15	1.753	37	1.687	59	1.671
16	1.746	38	1.686	60	1.671
17	1.740	39	1.685	70	1.667
18	1.734	40	1.684	80	1.664
19	1.729	41	1.683	90	1.662
20	1.725	42	1.682	100	1.660
21	1.721	43	1.681		
22	1.717	44	1.680		

The Chebychev Inequality is a procedure that may be used to establish an upper confidence limit for data that are not normally distributed. Persons evaluating the data should calculate an upper confidence limit for the background data by using the Chebychev Inequality equation (equation 12).

(12) $UCL = \bar{X} + k * s/(n^{1/2})$

Where: UCL = the upper confidence limit
k = the 95% Chebychev upper confidence value (calculated from equation. 12)
n = number of observations (values) in the background data set
$\bar{X}$ = sample mean
s = sample standard deviation

(13) $k = (1/(1-cl/100) -1)^{1/2}$

Where cl = confidence limit (in this case 95 since we are using an upper 95% confidence interval)

The upper confidence limit of the background mean derived by the Chebychev equation then may be compared with the mean of the monthly vapor reading that triggered the release investigation plus the additional vapor samples collected as part of the release investigation. If the mean of the monthly result plus the investigative data exceeds the upper confidence limit of the background mean, a release may be suspected. The tank owner/operator must report this result to DEQ and determine if a release has occurred.

Example 4. Determining if the mean of monthly plus confirmatory samples exceeds the background mean (Normally Distributed Data).

Background Vapor Concentrations in ppm (data are normally distributed)

100	53	350	1000	1200	49	350	600	650	400
850	90	700	245	1345	75	130	550	920	240

$\bar{X}$ = 494.9 ppm ~ 495 ppm
s = 401.6 ppm ~ 402 ppm
k at the 95% confidence limit is 2.396

the upper tolerance limit (TL) = X + k * s = 1459 ppm

Monthly vapor monitoring is initiated at the site and the vapor reading for the eighth month is 1900 ppm. The tank owner/operator takes four additional vapor measurements over the next week and these concentrations are: 1400, 750, 1600, and 500 ppm. The mean of the monthly reading that triggered the investigative monitoring and the additional vapor readings is 1230 ppm. The standard deviation of these vapor concentrations is 587 ppm.

Welch's t-test may be used to determine if the mean of this data is statistically greater than the background mean.

$t = (X_c - X_b) / [s_c^2 / n_c + s_b^2 / n_b\ ]^{1/2}$
$t = (1230 - 495) / [587^2 / 5 + 402^2 / 20\ ]^{1/2}$
$t = 735/[68914 + 8080]^{1/2}$
$t = 2.65$

$df = [s_c^2/n_c + s_b^2/n_b]^2 / [(s_c^2/n_c)^2/(n_c-1) + (s_b^2/n_b)^2/(n_b-1)]$
$df = [587^2 / 5 + 402^2 / 20\ ]^2\ / (587^2/5)^2/4 + (402^2/20)^2/19]$
$df = 5$

The critical statistic (t_c) for 5 degrees of freedom is 2.015 (Table 4).
t > tc, therefore, we reject the null hypothesis that there is no significant difference between the two means at the 95th percent confidence level. At the 95th percent confidence level, there is evidence that the mean concentration of the additional vapor monitoring data is statistically greater than the background mean.

Example 5. Determining if the mean of monthly plus confirmatory samples exceeds the background mean (data that are not normally distributed).

Background data:

0	0	25	33	45	0	12	0	0	95
32	34	56	89	0	9	0	12	34	15

n = 20 $\bar{X} = 24.6$ s = 28.7

Monthly monitoring and data from release investigation:
125 89 61 45 75

mean of monthly monitoring data and data from release investigation:
79

$UCL = \bar{X} + k * s/(n^{1/2})$

$k = (1/1\text{-cl}/100) - 1)^{1/2}$
$k = (1/(1-95/100) - 1)^{1/2}$
k = 4.359

$UCL = 24.6 + 4.359 * 28.7/(20)^{1/2}$
UCL = 52.6 ˜ 53

The mean of the monthly monitoring and data from the release investigation is greater than the upper confidence limit for the background data (79 > 53). A release should be suspected.

References

EPA. 1998. Draft EPA Guidance on Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities.

Lieberman, Gerald F. 1958. "Tables for One-Sided Statistical Tolerance Limits." Industrial Quality Control. Vol. XIV, No. 10.

Virginia Department of Environmental Quality (DEQ). 1998. Introduction to Groundwater/Soil Statistics.

DISCLAIMER

This document provides procedural guidance to the DEQ Storage Tank Program Staff. This document is guidance only. It does not establish or affect legal rights or obligations. It does not establish a binding norm and is not finally determinative of the issues addressed. Agency decisions in any particular case will be made by applying the State Water Control Law and the implementation regulations on the basis of site-specific facts.

USER'S GUIDE
STATISTICS - VAPOR MONITORING FOR RELEASE DETECTION

The purpose of this "user's guide" is to provide instructions for tank owners/operators who chose to use the spreadsheet developed by the DEQ Storage Tank Program entitled: : *Statistics - Vapor Monitoring for Release Detection*.

An upper tolerance limit, the concentration at which a release may be suspected needs to be derived for each vapor monitoring well at a site or facility. The steps listed below need to be followed to derive an upper tolerance limit for a single vapor monitoring well.

1. Open the MS. Excel Spreadsheet entitled: *Statistics - Vapor Monitoring for Release Detection*. Make sure that the macro's are enabled when the file is opened.

2. Go to cell *A72* and clear all data by pressing the "**clear**" button.

3. Enter at least 20 vapor readings, the "background data set" for an individual vapor monitoring well into cells *A6* through *A55*. Data that are below the instrument detection limit should be entered as 1/2 of the detection limit (e.g. if the detection limit is 1, enter a value of .5). Enter only numeric values, do not enter units (e.g. parts per million; ppm).

   NOTE At least 20 data values must be entered into cells A6 through A55 for the spreadsheet to perform the necessary calculations.

4. After entering the background data, check cells *B6* through *B55* to determine if any of the background vapor concentrations is an outlier. If none of the background vapor readings is an outlier, go to step 7. If any of the cells indicates that the corresponding vapor reading is an outlier, remove the vapor reading from the background data set.

   NOTE: An outlier is a data value, in this case a vapor concentration, that is statistically improbable when the data distribution is evaluated. Outliers often are indicative of sample collection errors.

5. Collect one additional vapor reading for each outlier found in step 4.

6. Record the vapor reading(s) replacing the outlier(s) in the appropriate cell(s) in column A and go to step 4.

7. Go to cell *C5* and press the "**sort 1**" button.

8. Go to cell *D5* and press the "**sort 2**" button.

9. The value listed in cell *D11* is the <u>upper tolerance limit</u> for the background vapor concentrations at the site. The upper tolerance limit is the action level for the vapor monitoring well. Save the file (it is recommended that you change the name of the file to reflect the well number of the pertinent vapor monitoring point).

10. Close the file. If there is another vapor monitoring well for which an upper tolerance limit has not been determined, return to step 1 and determine the upper tolerance limit for that well. If upper tolerance limits have been determined for each well, monthly vapor monitoring for release detection may be initiated.

11. Perform release detection by monitoring vapor concentrations in each vapor monitoring well. This is required on a monthly basis.

12. Compare the monthly vapor monitoring concentration in each well with the upper tolerance limit for that well (cell *D11*). If the vapor concentration is below the upper tolerance limit, return to step 11. If the vapor concentration for any well exceeds the upper tolerance limit for that well, go to step 13.

13. Contact the DEQ Regional Office within 24 hours after discovering that a vapor monitoring result exceeded the upper tolerance limit for that particular well. The DEQ Regional Office staff will provide further instructions and a deadline for performing the tests needed to further evaluate whether the vapor monitoring data suggest a release has occurred.

14. Open the MS Excel Spreadsheet file for the vapor monitoring well in which the monthly vapor reading exceeded the upper tolerance limit. Make sure that the macro's are enabled when the file is opened.

15. Enter the vapor reading that exceeded the upper tolerance limit into cell *E6*.

16. Collect four (4) to nine (9) additional vapor readings from the vapor monitoring well over the next week and enter these values into cells *E7* through *E17*.

17. Go to cell *D14* and press "**copy A**".

18. Go to cell *D16* and press "**copy B**".

19. Report the results of the statistical test listed in cell *D65* to the Regional Office within the time specified by the Regional Office (see step 13)